FLUIDEZ NAS COMUNICAÇÕES

ANACOM AUTORIDADE NACIONAL DE COMUNICAÇÕES

Postal traffic in Portugal
Applying time series modeling

27TH CONFERENCE ON POSTAL AND DELIVERY ECONOMICS
24 May 2019
Dublin

Filipa Silva
Carla Machado
ANACOM | Portugal

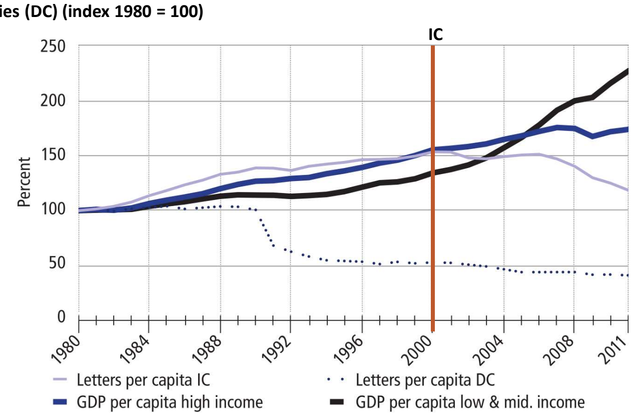
Table of contents

1. **Brief context**
2. **Data source and samples**
3. **Postal traffic evolution and determinants**
4. **Methodological framework**
5. **Estimation, goodness-of-fit and forecast**
6. **Conclusions**

1. Brief context

Past: GDP - the main driver

Evolution of letter-post volume per capita (domestic and international combined) and GDP per capita in industrialized countries (IC) and developing countries (DC) (index 1980 = 100)



Source: UPU Postal Statistics

Present: Digitalization rules

'Postal Economic Outlook 2018 – lastest trends in an evolving sector'. UPU (2018)

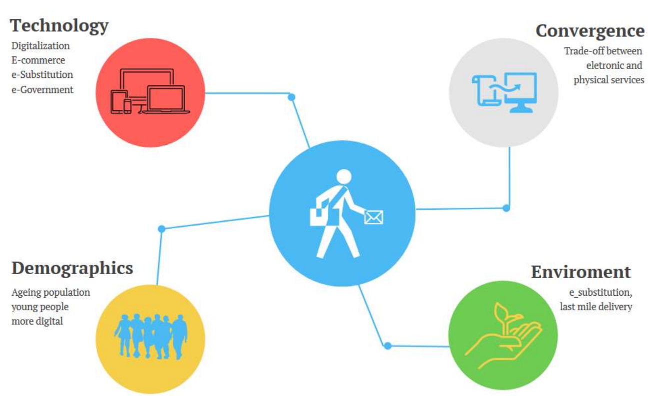
Demand for postal service

Internet affects: consumers behaviours and production and distribution of goods and services

The authors, based in information from UPU (2018).

Present: Digitalization rules

'Developments in the postal sector and implications for regulation'. ERGP (2019)



The authors, based in information from ERGP (2019).

2. Data source and samples

Data source

Source: Data from ANACOM, collected from postal services providers

Traffic (domestic and outgoing international), by type of item:

Correspondence
Letters, editorial mail and direct mail

Parcels

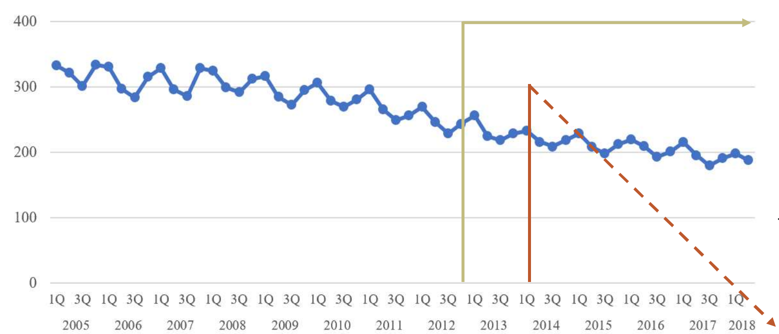
Time span: quarterly data, from

1Q 2005 → **2Q 2018**

3. Traffic evolution and determinants

Postal traffic (domestic and outgoing international)

Figure 3 – Evolution of the total postal traffic (domestic and outgoing international), Portugal



Law no. 17/2012, of 26 April: (transposition of the European Directive)

Revised definitions of postal services

Removed the maximum weight limit for postal parcels (previously, 20 Kgs)

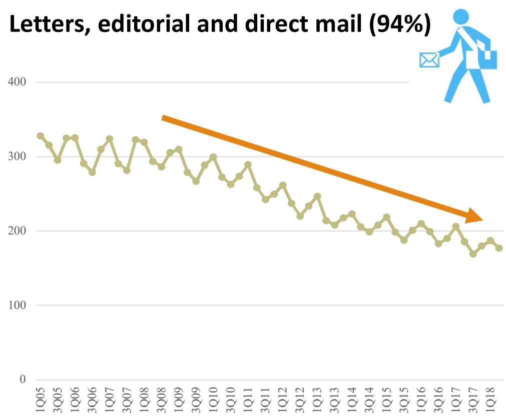
2014
New object types reclassification reported by some postal operators
(affects parcels, mostly)

Unit: Millions of objects.

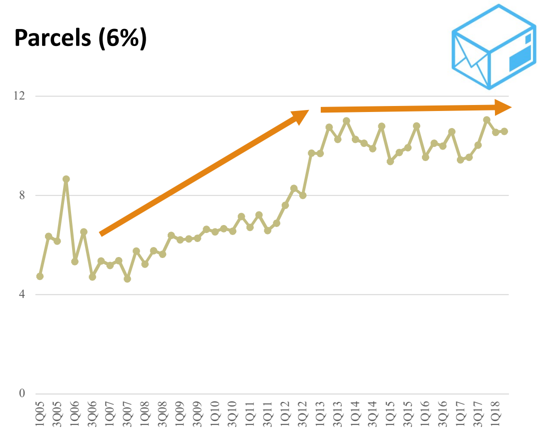
Source: Authors, with data from ANACOM.

… by types of items

Letters, editorial and direct mail (94%)



Parcels (6%)



Unit: Millions of objects.
Source: Authors, with data from ANACOM.

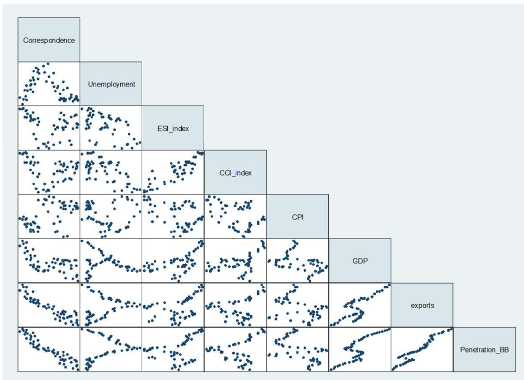
Explanatory variables

Table 1 – Variables considered that may affect postal traffic

Name	Description	Source
Unemployment	Total Unemployment	INE, Portugal
ESI_Index	Economic sentiment indicator	INE, Portugal
CCI_Index	Consumer confidence indicator	INE, Portugal
CPI	Consumer price index (12-month average growth rate)	INE, Portugal
GDP	Gross domestic product	INE, Portugal
Exports	Exports of goods and services	INE, Portugal
Penetration_BB	Fixed broadband accesses, per 100 inhabitants	ANACOM

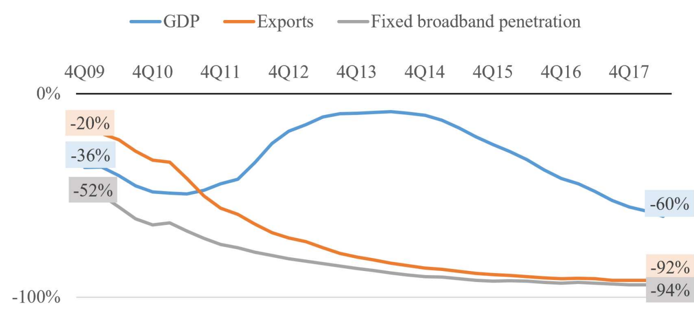
Correspondence

Figure 6 - Correlations between correspondence traffic and exogenous variables (1Q2005 – 2Q2018)



Sources: INE (Unemployment – total Unemployment; ESI_Index - Economic sentiment indicator; CCI_Index - Consumer confidence indicator; CPI - Consumer price index (12-month average growth rate); GDP – Gross domestic product; Exports – Exports of goods and services) and ANACOM (Penetration_BB – Fixed broadband accesses per 100 inhabitants).

Figure 7 – Correlation between correspondence traffic (letters, editorial mail and direct mail) and GDP, exports and fixed broadband penetration, between the 1Q2005 and the date mentioned in horizontal line, Portugal

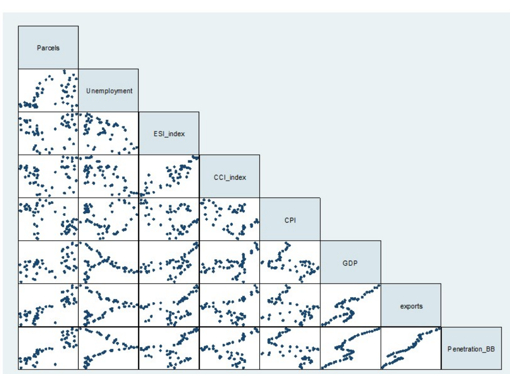


Unit: %.

Source: Authors, with data from ANACOM.

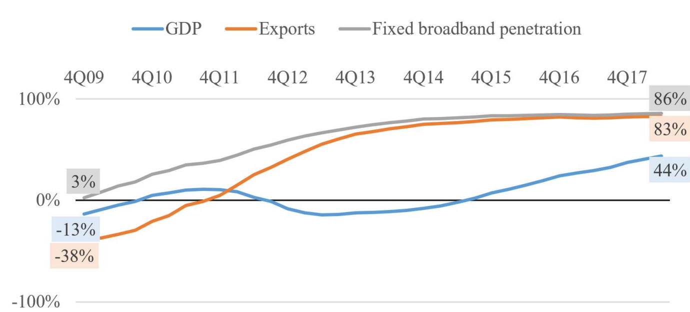
Parcels

Figure 9 - Correlations between parcels traffic and exogenous variables (1Q2005 – 2Q2018)



Sources: INE (Unemployment –total Unemployment; ESI_Index - Economic sentiment indicator; CCI_Index - Consumer confidence indicator; CPI - Consumer price index (12-month average growth rate); GDP – Gross domestic product; Exports – Exports of goods and services) and ANACOM (Penetration_BB – Accesses fixed broadband per 100 inhabitants).

Figure 10 – Correlation between parcels traffic and GDP and Internet penetration, between the 1Q05 and the date mentioned in horizontal line, Portugal



Unit: %

Source: Authors, with data from ANACOM

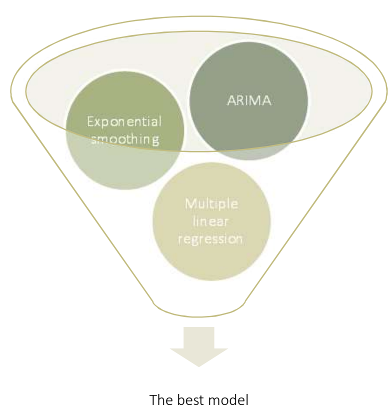
4. Methodological framework

Methodological approaches

(1) ARIMA models

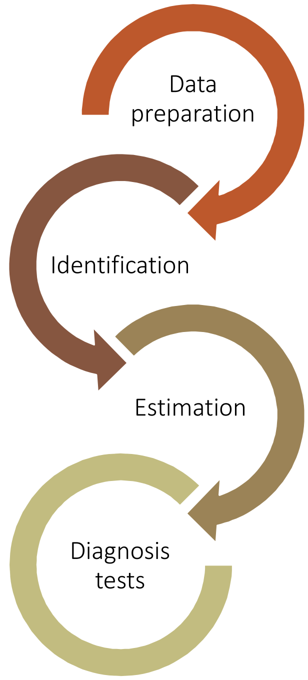
(2) Decomposition models:
exponential smoothing

(3) Multiple linear regressions



ARIMA MODELS

iterative process



Step 1

Transformations the data

Stationarity

Step 2

Orders of the ARIMA model: ACF and PACF

ARIMA(p,d,q) or SARIMA (p,d,q)(P,D,Q)s

Step 3

The estimation of the parameters can be done by the conditional least squares and maximum likelihood.

Step 4

Analyzing the residuals of the estimation.

Different ARIMAs can be estimated.

Compare the results after estimation

Error measures in the estimation period

- Root mean square error (RMSE).
- Bayesian information criterion (BIC)
- Akaike's information criterion (AIC)

Residual diagnostics and goodness-of-fit tests

- Residual autocorrelation and cross correlation plots
- Durbin-Watson statistic (serial correlation test);
- Non-normality test: skewness / kurtosis; Shapiro-Wilk W statistic
- Heteroscedasticity test: Breusch-Pagan test / White's test
- Misspecification test: Ramsey RESET test

Error measures in the validation period

Out-of-sample

- Root mean squared forecast error (RMSFE)

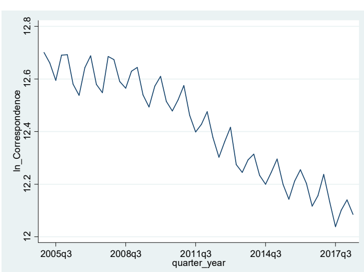
Qualitative considerations

- The appearance of forecast plots, intuitive reasonableness of the coefficients and the simplicity of the model.

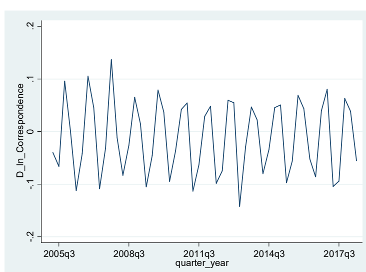
5. Estimation, goodness-of-fit and forecast

5.1. Correspondence traffic | Time series

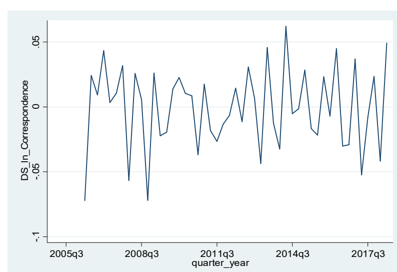
Log transformation
ln_correspondence



1st differences
D_ln_correspondence



Seasonal differences (period 4)
DS_ln_correspondence



Time series: From 1Q2005 to 2Q2018 (54 observations)

5.1. Correspondence traffic | Estimation

ARIMA MODELS

(1.1) SARIMA (p,d,q) (P,D,Q)s:

$\Phi(B^s)\phi(B)\nabla_S^D\nabla^d y_t = \Theta(B^s)\theta(B)\varepsilon_t$

SARIMA (0,1,1) $(0,1,1)_4$

(1.2) SARIMAX: **with exogenous variable**

$\Phi(B^s)\phi(B)\nabla_S^D\nabla^d y_t = \Psi(B)X_t + \Theta(B^s)\theta(B)\varepsilon_t$

SARIMAX (2,1,1) $(0,1,0)_4$ | X_t =ln_GDP

DECOMPOSITION MODELS

(1.3) Holt-Winters' Multiplicative

$$\begin{aligned}
\text{level} \quad L_t &= \alpha\frac{y_t}{S_{t-s}} + (1-\alpha)(L_{t-1}+b_{t-1}); \\
\text{trend} \quad b_t &= \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}, \\
\text{seasonal} \quad S_t &= \gamma\frac{y_t}{L_t} + (1-\gamma)S_{t-s} \\
\text{forecast} \quad F_{t+k} &= (L_t + kb_t)S_{t+k-s},
\end{aligned}$$

MULTIPLE LINEAR REGRESSION

$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_p X_{pt} + \varepsilon_t$

(1.4) Without exogenous variable:

- Linear trend: t
- Seasonal dummies: Q1, Q2, Q3
- Structural breaks: D4Q2007; D4Q2011 - linear effect

(1.5) With exogenous variable:

- Exogenous variable: ln(penetrationBB)
- Seasonal dummies: Q1, Q2, Q3
- Structural breaks: D4Q2011
- Interaction regressors

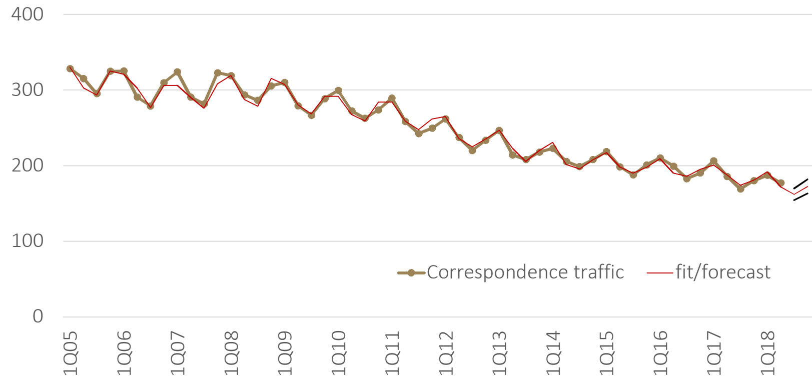
5.1. Correspondence traffic | Goodness-of-fit evaluation

	(1.1) SARIMA	**(1.2) SARIMAX**	**(1.3) Holt-Winters' Multiplicative**	**(1.4) MLR**	**(1.5) MLR with ln(PenetrationBB)**
RMSE	25,54	25,25	23,01	23,26	26,68
AIC	-213,72	-217,62	(*)	-243,32	-222,63
BIC	-206,16	-203,69	(*)	-229,39	-208,71
Out-of-sample (RMSFE)**	24,46	25,16	20,31	20,95	25,25

Notes: (*) not comparable indicators;

(**) Comparison between predicted values to real values of the time series. The set forecasts to start of 1st quarter 2018 to the end of 2nd quarter 2018.

5.1. Correspondence traffic | Forecast



Correspondence traffic decrease:
- 7% in 3Q2018
- 5% in 4Q2018

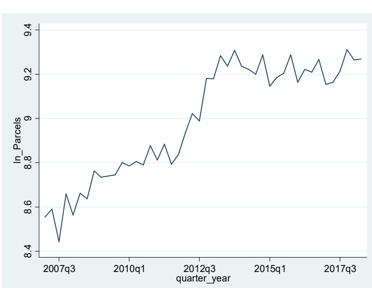
(both from the previous year)

Absolute error around 5 p.p.

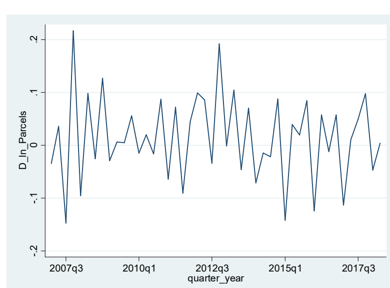
Unit: Millions of objects.
Source: Authors, with data from ANACOM.
Note: Lower confidence limits (LCL) and upper confidence limits (UCL) for 95% confidence interval.

5.2. Parcels traffic | Time series

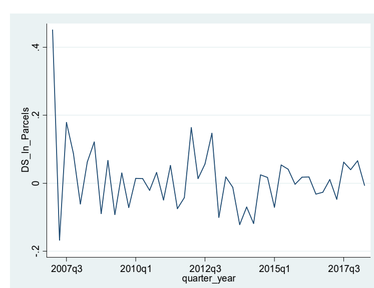
Log transformation
ln_parcels



1st differences
D_ln_parcels



Seasonal differences (period 4)
DS_ln_parcels



Time series: From 1Q2007 to 2Q2018 (46 observations)

5.2. Parcels traffic | Estimation

ARIMA MODELS

(1.1) ARIMA (p,d,q):

$\phi(B)\nabla^d y_t = \theta(B)\varepsilon_t$

ARIMA (4,1,0)

(1.2) ARIMAX: with exogenous variable

$\Phi(B^s)\phi(B)\nabla_S^D\nabla^d y_t = \Psi(B)X_t + \Theta(B^s)\theta(B)\varepsilon_t$

ARIMAX (0,1,0) $(1,0,1)_4$ | X_t =lnPenetrationBB

DECOMPOSITION MODELS

(1.3) Holt-Winters' Additive

$$\begin{aligned}
\text{level} \quad L_t &= \alpha(y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1}); \\
\text{trend} \quad b_t &= \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}, \\
\text{seasonal} \quad S_t &= \gamma(y_t - L_t) + (1-\gamma)S_{t-s} \\
\text{forecast} \quad F_{t+k} &= L_t + kb_t + S_{t+k-s},
\end{aligned}$$

MULTIPLE LINEAR REGRESSION

$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_p X_{pt} + \varepsilon_t$

(1.4) Without exogenous variable:
- Linear trend: t
- Seasonal dummies: Q2, Q3, Q4
- Structural breaks: D4Q2012; D1Q2014
- Interaction regressors

(1.5) With exogenous variable:
- Exogenous variable: ln(penetrationBB)
- Seasonal dummies: Q4
- Structural breaks: D4Q2012; D1Q2014
- Interaction regressors

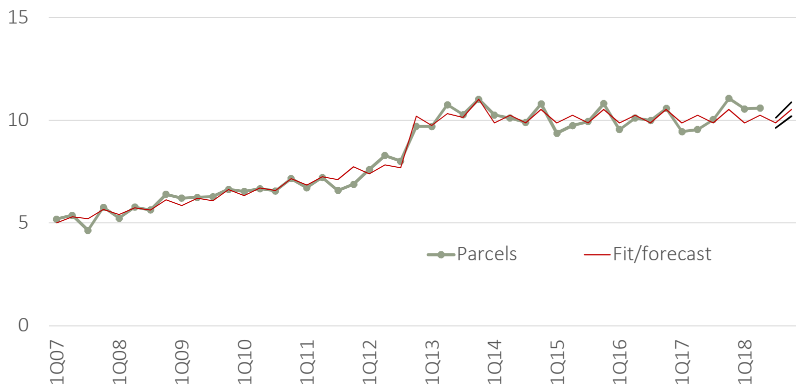
5.2. Parcels traffic | Goodness-of-fit evaluation

	(1.1) ARIMA	(1.2) ARIMAX	(1.3) Holt-Winters' Additive	(1.4) MLR	(1.5) MLR with ln(PenetrationBB)
RMSE	61,69	57,70	50,27	**41,32**	48,06
AIC	-116,9	-121,8	(*)	-150,61	-138,71
BIC	-109,7	-112,8	(*)	-139,64	-129,57
Out-of-sample (RMSFE)**	**20,55**	27,38	59,24	30,75	36,59

Notes: (*) not comparable indicators;

(**) Comparison between predicted values to real values of the time series. The set forecasts to start of 1st quarter 2018 to the end of 2nd quarter 2018.

5.2. Parcels traffic | Forecast



Unit: Millions of objects.
Source: Authors, with data from ANACOM.
Note: Lower confidence limits (LCL) and uppr confidence limits (UCL) for 95% confidence interval.

Parcels traffic:

Null variation (0%) from the previous year

Around 10 million objects

Absolute error around 3 p.p. (both from the previous year)

6. Conclusion

1

IT sector has two different effects on postal sector – showed by Portuguese data

- Digitalization (e-government, e-substitution and e-invoice) has a negative impact on the correspondence postal traffic.
- E-commerce helped the parcels traffic to grow, due to the delivery of physical product brought through the Internet.

2

GDP is gaining importance again to explain postal traffic

- GDP lost its force to explain these series, mainly due the financial crisis. In recent years, it is gaining importance again, especially in the correspondence postal traffic.
- However, it has a negative effect on correspondence, in contrast with what happened in the past.

3

The decrease of the correspondence traffic
is not expected to slow down soon, in Portugal

- Between 1Q2005 and 2Q2018, the best fitted model is Multiplicative Holt-Winters
- Forecasts show a decrease of correspondence traffic:
    - around 7% in 3Q2018 (from the previous year)
    - around 5% in 4Q2018 (from the previous year)
    - with an absolute error around 5 percentual points

4

The forecast parcels traffic shows a stabilization in Portugal, but may increase in the future, with the influence of other variables

- Between 1Q2007 and 2Q2018, the best fitted model is the Multiple Linear Regression, with a trend, seasonality dummies, a structural break dummy and no exogenous variables
- Forecast shows a stabilization around 10 million objects, with an absolute error around 3 percentual points (from the previous year)

FLUIDEZ NAS COMUNICAÇÕES

ANACOM AUTORIDADE NACIONAL DE COMUNICAÇÕES

Thank you!